Title page

Multisensory Signals Inhibit Pupillary Light Reflex: Evidence from Pupil Oscillation

Xiangyong Yuan[a,b,c,*], Yuhui Cheng[a,b,c], Yi Jiang[a,b,c,d,*]

a. State Key Laboratory of Brain and Cognitive Science, CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China

b. Department of Psychology, University of Chinese Academy of Sciences, Beijing, 100049, China

c. Chinese Institute for Brain Research, Beijing, 102206, China

d. Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, 230088, China

Short title: Multisensory inhibition of pupillary light reflex

* To whom correspondence should be addressed. Email: yuanxy@psych.ac.cn or yijiang@psych.ac.cn

Abstract

Multisensory integration, which enhances the stimulus saliency at the early stage of processing hierarchy, is recently shown to produce a larger pupil size than its unisensory constituents. Theoretically, any modulation on pupil size ought to be associated with the sympathetic and parasympathetic pathways that are sensitive to lights. But it remains poorly understood how pupillary light reflex is changed in a multisensory context. The present study evoked an oscillation of pupillary light reflex by periodically changing the luminance of a visual stimulus at 1.25 Hz. It was found that such induced pupil oscillation was substantially attenuated when the bright but not the dark phase of the visual flicker was periodically and synchronously presented with a burst of tones. This inhibition effect persisted when the visual flicker was task-irrelevant and out of attentional focus, but disappeared when the visual flicker was moved from the central field to the periphery. These findings not only offer a comprehensive characterization of the multisensory impact on pupil response to lightness, but also provide valuable clues to the individual contributions of the sympathetic and parasympathetic pathways to multisensory modulation of pupil size.

Keywords: multisensory; pupil size; pupillary light reflex; oscillation; stimulus eccentricity; task relevance

1 Introduction

Combining various information from distinct sensory modalities is beneficial for interaction with the environment. For instance, many have shown that multisensory integration facilitates detection, discrimination and search (Leo, Bertini, di Pellegrino, & Làdavas, 2008; Noesselt, Bergmann, Hake, Heinze, & Fendrich, 2008; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008), amplifies the activation of sensory cortical areas (Kayser, Philiastides, & Kayser, 2017; Lewis & Noppeney, 2010; Noesselt et al., 2010; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011; Werner & Noppeney, 2010, 2011) and subcortical nucleus (most importantly, the superior colliculus, see Stein & Stanford, 2008; Stein, Stanford, & Rowland, 2020). All these evidence reflects an enhancement of stimulus saliency by multisensory integration at an early processing stage. Since our pupil size is sensitive to salient stimulus, with larger pupil size corresponding to stimulus with higher saliency (e.g., objectively high contrast, or subjectively easy-to-notice) irrespective of its modality (Liao, Kidani, Yoneya, Kashino, & Furukawa, 2016; Wang, Boehnke, Itti, & Munoz, 2014; Wang & Munoz, 2014), it is assumed that multisensory signals could dilate the pupil size to a larger degree than its unisensory constituents.

The breakthrough came from a study on rhesus monkey, which found that concurrently presented flash and beep in periphery elicit a transient pupil dilation, equaling the linear summation of the pupil size when they were presented in isolation (Wang et al., 2014). This finding was later replicated on humans by two independent studies, which further indicate in a detection task that the larger the pupil size, the faster the saccadic or manual response to the audiovisual stimuli (Rigato, Rieger, & Romei, 2016; Wang, Blohm, Huang,

Boehnke, & Munoz, 2017). Moreover, it is shown that the enlarged pupil size when visual stimuli are presented in the central field in combination with auditory stimuli exceeds the linear summation of the pupil size obtained in each modality (Rigato et al., 2016, but see Van der Stoep, Van der Smagt, Notaro, Spock, & Naber, 2021). As acknowledged, the pupil size is controlled by two antagonistic pathway, the sympathetic pathway that enlarges the pupil size and the parasympathetic pathway that constricts the pupil size (Eckstein, Guerra-Carrillo, Miller Singley, & Bunge, 2017; Joshi & Gold, 2020; Larsen & Waters, 2018; Wang & Munoz, 2015). Therefore, the pupil dilation induced by multisensory integration may reflect either an increased sympathetic activation, or a decreased parasympathetic activation, or their combination (refer to the discussion of Wang et al., 2014 for more details).

Notably, these two pathways are sensitive to ambient luminance. Pupil constriction to brightness (or pupillary light reflex) is mainly driven by the parasympathetic activation, while pupil dilation to darkness is mainly driven by the sympathetic activation[1] (Joshi & Gold, 2020). Investigations on how pupillary responses to different light levels are modulated in a multisensory context can provide insightful clues to the individual contributions of the two pathways to such modulation. It has already been shown that the onset latency of pupil dilation evoked by stimulus saliency could be as early as that of pupillary light reflex, which suggests that the initial component of the transient pupil dilation induced by higher visual contrast is probably a result of the inhibition of the parasympathetic activation (Wang & Munoz, 2014). It is thus presumed that multisensory

[1] Of note, this is a straightforward and simplified statement and both the parasympathetic and sympathetic pathways may engage in modulation of pupil response to lights (ref to Box 1 in Joshi & Gold, 2020).

signals, if enhance stimulus saliency, are able to specifically inhibit the parasympathetic activation in a very short time, which may in turn attenuate the pupillary light reflex transiently. However, this hypothesis that multisensory signals could inhibit pupillary light reflex has rarely been empirically tested.

To probe this issue, the present study, following the pupil frequency tagging method (Naber, Alvarez, & Nakayama, 2013), periodically presented a simple, emotionally neutral stimulus and modulated its luminance at 1.25 Hz to elicit an oscillation of pupil size. In a series of four experiments, we presented a tone periodically at the same frequency with the repeated visual stimulus and manipulated the temporal congruency between the tone pulses and the bright phase of the visual flicker. Using this method, when the tone synchronizes with the light phase, the amplitude of this pupil oscillation can be employed as a quantitative measure of the multisensory impact on the pupillary light reflex. In contrast, when the tone synchronizes with the dark phase, the oscillatory amplitude quantifies the multisensory impact on the dark reflex (or on the relaxation of pupillary light reflex). We examined whether this pupil oscillation is attenuated by the tone pulses synchronous with the bright phase (Experiments 1 and 2) and further delineated the respective roles of stimulus eccentricity and task relevance in the multisensory inhibition of pupillary light reflex (Experiments 3 and 4).

2 Experiment 1

Experiment 1 examined whether multisensory inputs inhibit pupillary light reflex. The visual flickering stimulus, which changes its luminance periodically, would induce a dynamic change of pupil size, or in other words an oscillation of pupil size. If multisensory inputs

inhibit light reflex, the pupil oscillation would fluctuate in a smaller range (i.e., a smaller oscillatory amplitude) when the auditory stimuli are temporally congruent with the bright phase of the visual flicker despite the actual luminance remains constant.

2.1 Methods

2.1.1 Participants

Sixteen participants were recruited in Experiment 1 (8 females; mean age: 21.9 ± 2.7 years). All participants had normal or corrected-to-normal vision and normal hearing, and were naïve to the purpose of the experiment. They provided informed written consent before experiment and were paid for their participation after experiment. The study was approved by the institutional review board of the Institute of Psychology, Chinese Academy Sciences (H18029), and adhered to the tenets of the Declaration of Helsinki.

2.1.2 Stimuli and apparatus

A pioneer study has revealed that pupil oscillation is evoked by visual stimuli flickering at a frequency below ~ 3 Hz (Naber et al., 2013). In Experiment 1, a disc presented in the central field (radius: 1.61 degree of visual angle), which flickered between brightness (22.56 cd/m^2) and darkness (15.15 cd/m^2) at 1.25 Hz, was used as the visual stimuli (Fig.1a). The auditory stimulus was a tone (carrier frequency: 700 Hz; sample rate: 44100 Hz) with a duration of 0.4 secs, played binaurally through headphones (Sennheiser HD 201). The loudness of the tone was set at a comfortable sound level throughout the experiment (~ 60 dB (A)) and kept constant for all participants.

The experiment was conducted in a dim, sound-attenuated room. Participants sat comfortably at a viewing distance of about 60 cm from the screen (refresh rate: 60 Hz,

resolution: 1920 × 1080). The luminance of the gray background was 18.67 cd/m^2. All stimuli were generated by Matlab (The MathWorks Inc.) and presented using Psychtoolbox (Brainard, 1997; Pelli, 1997). Pupil size and eye position of the left eye were recorded using a video-based iView X Hi-Speed system (SMI, Berlin, Germany) at 500 Hz. Participants put their heads on a chin-rest and were told to minimize head movements during the recording period. The recorded pupil size was analyzed and reported in arbitrary unit (a.u.) without transformed into actual unit (mm), as the relative change of the pupil size was of our main interest. In general, a pupil size of ~33 a.u. corresponded to a pupil size of 5 mm in the present study.

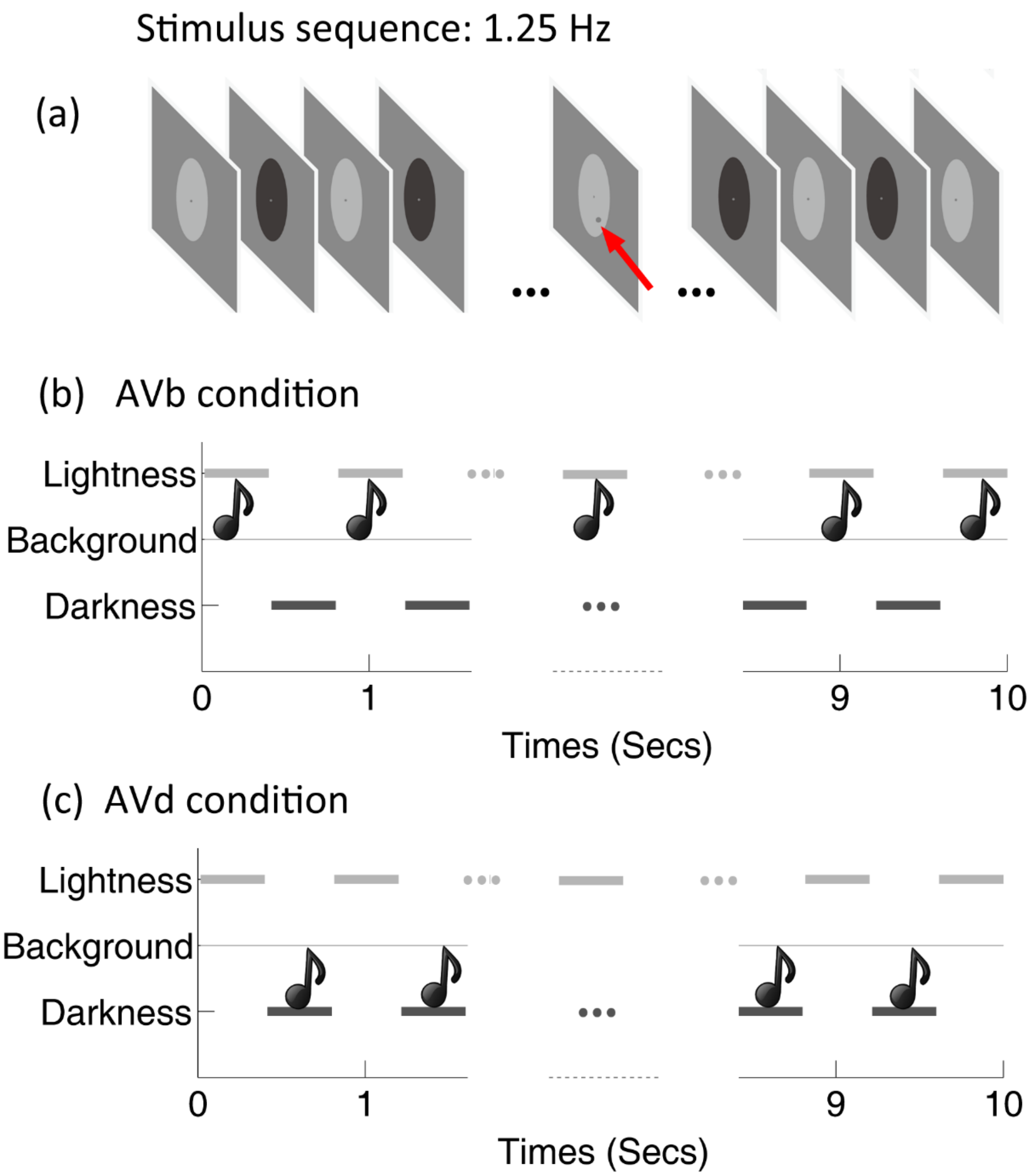

Figure 1. Stimulus and an exemplar trial. (a) The luminance of the disc modulated at 1.25 Hz. The red arrow points out the oddball dot that participants had to count. (b)(c) The tone is synchronized with the bright phase of the disc in the AVbright condition (AVb), while synchronized with dark phase of the disc in the AVdark condition (AVd).

2.1.3 Procedures

In each trial, the fixation (a small dot, diameter: 0.16 º) was first presented as a warning

signal to inform the participants that they should fixate at this position, prepare for the appearance of the visual stimuli and avoid eye blinks. After a random duration of 1.5 – 2 secs, the flickering disc was presented for 10 secs (Fig.1a). To maintain participants' attention on the disc, they were required to complete an oddball counting task, in which small dots (diameter: 0.27 º) flashed for 0.05 secs at random positions of the disc, and participants count how many times they saw the oddballs. There were a total of 0 – 3 oddballs, randomly determined for each trial and never being presented at the same time. The oddball, if presented at the bright phase of the disc, had an equal luminance with the dark phase of the disc, and vice versa. After inputting their answers, participants could relax their eyes for a while and then press the SPACE key to initiate the next trial.

There were four conditions in Experiment 1. In the visual-only condition (V-only), the disc was presented silently. In the auditory-only condition (A-only), the tone was periodically presented at 1.25 Hz, but the luminance of the disc remained constant, either bright or dark. The tone was synchronized with the bright phase of disc in the AVbright condition (AVb), while synchronized with the dark phase of the disc in the AVdark condition (AVd; Fig.1b and 1c). There were 64 trials in total, divided into 4 blocks. In each block, each condition was repeated 4 times. A 5-point standard calibration of the eye position was routinely conducted before the first block and third block, but if necessary, before any other blocks.

2.1.4 Data analysis

The accuracy of the oddball counting task was calculated as the number of trials with correct answers dividing by the total number of trials. The raw pupil diameter in each trial

was visually inspected, and trials with blinks more than three times and other artifacts were excluded (2.1 trials excluded on average). For the remaining trials, data points where the eye position deviated 3 SDs of the mean, the pupil diameter deviated 3 SDs of the mean, or dropped largely due to blinks or blink-like artifacts (i.e., the recording system failed to detect the corneal reflex but the pupil diameter still showed a blink-like shrink) were linearly interpolated. The artifact-free pupil diameter was then downsampled by averaging the data points in every 0.05 sec non-overlapping window, and detrended to minimize slow drift. To quantify pupil oscillation, fast Fourier Transform (FFT) was conducted for each trial, wherein the first second was excluded to remove the transient response to stimulus onset (Naber et al., 2013). The amplitude of pupil oscillation was calculated as the modulus of the FFT complex coefficients and averaged across trials for each condition. Finally, the amplitude spectra were normalized by subtracting the amplitude averaged across the neighboring four frequency points (within ± 0.156 Hz) from the amplitude at each frequency point.

2.1.5 Statistics

To evaluate whether the pupil size oscillated at 1.25 Hz, we performed one-sample t-tests on the normalized amplitude at 1.25 Hz for each condition, respectively. The normalized amplitude, if significantly larger than zero, indicates a robust pupil oscillation at that condition. In the next, we compared the normalized amplitude between conditions that observed significant pupil oscillation, using paired-sample t tests, to examine how multisensory signals modulate pupil oscillation. The reported p values were Bonferroni corrected for multiple comparisons if not specifically mentioned. In addition, we computed

the JZS Bayesian factor (BF_{10}, H1 versus H0) using a matlab toolbox developed by Bart Krekelberg, retrieved from GitHub (https://www.github.com/klabhub/bayesFactor). BF_{10} assesses the relative evidence for H1 over H0. A BF_{10} larger than 3 provides substantial evidence for H1, while a BF_{10} smaller than 1/3 provides substantial evidence for H0 (Dienes, 2014).

2.2 Results and discussion

The accuracy of the oddball counting task approached 100 % in all conditions (V-only: 0.98 ± 0.04; A-only: 0.97 ± 0.06; AVb: 0.99 ± 0.02; AVd: 0.98 ± 0.04), indicating that participants had focused their attention on the central flicker during eye recording. As seen in Fig. 2a and 2b, the pupil size oscillated during the presentation of the flicker in all except the A-only condition. One-sample t-tests confirmed the observation that the normalized amplitude of pupil oscillation at 1.25 Hz was significantly greater than zero in the V-only, the AVb and the AVd conditions ($ts > 9$, $ps < 4^{e-7}$, $BF_{10} > 1^{e+5}$), but not in the A-only condition ($t_{15} = 0.002$, $p > 0.9$, $BF_{10} = 0.255$; Fig. 2c and 2d). Therefore, the oscillatory amplitude in the A-only condition was excluded from the following comparisons when examining the effect of audiovisual impact on the pupil oscillation. As shown in Fig. 2d, paired-sample t-tests revealed that the strength of pupil oscillation significantly decreased when the tones were temporally congruent with the bright phase of the visual stimuli, relative to the visual stimuli presented alone (V-only vs AVb: $t_{15} = 3.032$, $p = 0.025$, $BF_{10} = 6.313$). No other significant effects were found (AVd vs AVb: $t_{15} = 1.475$, $p = 0.483$, $BF_{10} = 0.632$; V-only vs AVd: $t_{15} = 0.111$, $p > 0.9$, $BF_{10} = 0.257$).

Experiment 1 showed that pupil oscillation was induced by luminance modulation of visual stimulus, in accordance with previous findings (Naber et al., 2013). More importantly, it indicated that the pupillary light reflex was suppressed in a multisensory context. By contrast, when the tones were synchronized with the dark phase of the visual flicker, the pupillary responses were not significantly changed. Therefore, the relatively fast pupil frequency tagging method in Experiment 1 specifically captured a multisensory inhibition on pupillary light reflex with virtually no impact on the dark reflex (or relaxation from pupillary light reflex; ref to the *General Discussion* section for the possible account of this finding). In order to replicate the results, we conducted Experiment 2. Instead of luminance modulation, we periodically flashed a disc which was either brighter (Experiment 2a) or darker (Experiment 2b) than the background, and played a tone synchronously at the onset time of the disc. Through this method, we could induce the pupil oscillation as in Experiment 1, and examined whether the tones had distinct impacts on the strength of pupil oscillations from Experiments 2a and 2b.

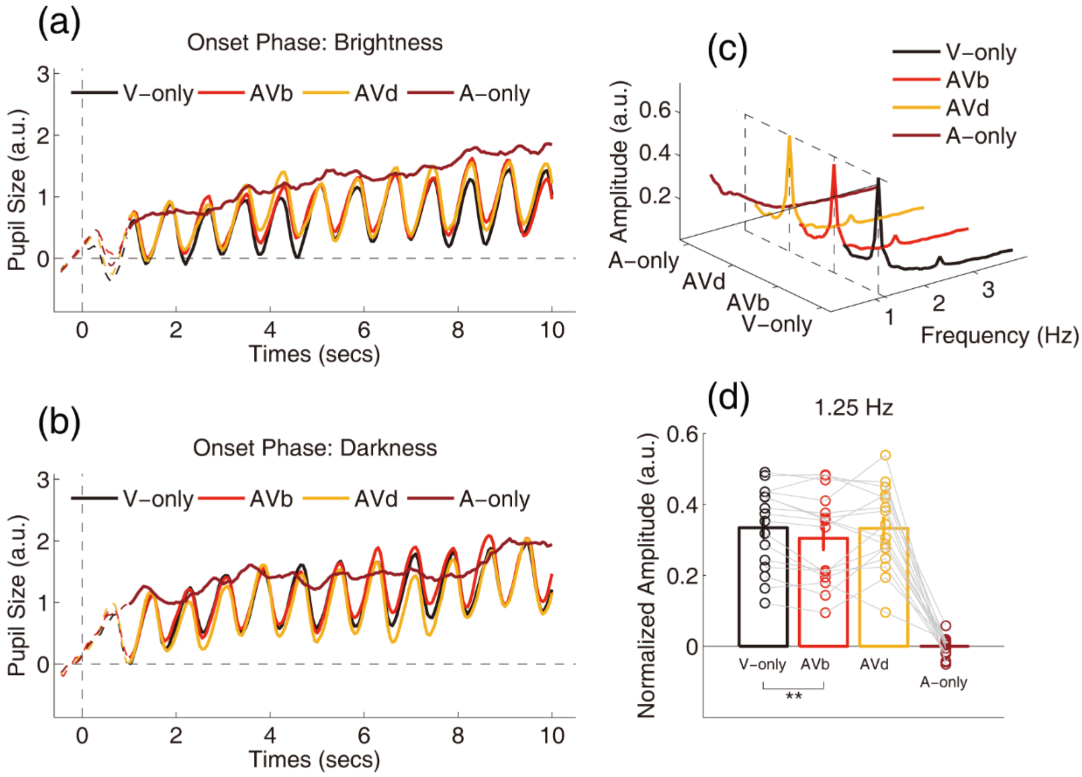

Figure 2. Results of Experiment 1. The baseline-corrected oscillation of pupil size when the disc started flickering from the bright phase (a) or the dark phase (b). The dashed color lines represent pupil size in the first second of the trial, which is excluded from FFT analysis. (c) The amplitude spectra after FFT. The dashed lines indicate the target frequency 1.25 Hz. (d) The normalized oscillatory amplitude at 1.25 Hz. Each circle represents the amplitude of pupil oscillation from one participant. The error bar indicates the standard error of mean. ** means $p < 0.01$, uncorrected. AVb represents the AVbright condition; AVd represents the AVdark condition.

3 Experiment 2

In Experiment 2, the visual stimulus was repeatedly presented against the background, with the tone pulses either synchronous with the stimulus or not. If Experiment 1's finding was robust, we expected that in Experiment 2a, where the visual stimulus was brighter

than the background, the pupil oscillation would be suppressed by synchronous tones, whereas in Experiment 2b, there was still not an increased pupil oscillation by synchronous tones when the visual stimulus was darker than the background.

3.1 Methods

3.1.1 Participants

Thirty-two new participants took part in Experiment 2, with 16 in Experiment 2a (12 females; mean age: 21.8 ± 2.5 years) and 16 in Experiment 2b (10 females; mean age: 21.2 ± 2.5 years).

3.1.2 Stimuli and apparatus

The luminance of the disc was always 32.40 cd/m^2 in Experiment 2a and 9.20 cd/m^2 in Experiment 2b. The duration of disc equaled 0.4 secs. The tone, and all other aspects were the same as Experiment 1.

3.1.3 Procedures

The main procedure of Experiment 2 was the same as that of Experiment 1, except that in each trial the disc flashed periodically at 1.25 Hz against the background to induce pupil oscillation. There were three conditions, V-only, AVb (in Experiment 2a) or AVd (in Experiment 2b), and AVbackground (AVbkg). In the V-only condition, the disc was presented alone. In the AVb or AVd condition, the tone and disc were simultaneously presented, while in the AVbkg condition, the tone was presented just when the disc disappeared. There were totally 48 trials, divided into 4 blocks. In each block, each condition was repeated 4 times.

3.1.4 Data analysis and statistics

The analysis and statistics were same as Experiment 1.

3.2 Results and discussion

Regardless of experiments and conditions, all participants performed well in the oddball counting task (V-only: 0.94 ± 0.06; AVb: 0.97 ± 0.04; AVbkg: 0.98 ± 0.03 in Experiment 2a, and V-only: 0.95 ± 0.04; AVd: 0.96 ± 0.06; AVbkg: 0.96 ± 0.04 in Experiment 2b). Apparent pupil oscillation was observed in all conditions of Experiment 2 (Fig. 3a and 3b, $ts > 7$, $ps < 4^{e-5}$, $BF_{10} > 5^{e+3}$; the pupil oscillation in each condition was drawn in Supplementary Fig. 1). The results of Experiment 2a replicated Experiment 1. The amplitude of pupil oscillation decreased when the tone was synchronized with the disc with a brighter luminance (Fig. 3a), compared with the disc were presented alone (V-only vs AVb: $t_{15} = 3.766$, $p = 0.006$, $BF_{10} = 22.385$) or accompanied by an asynchronous tone (AVbkg vs AVb, $t_{15} = 3.192$, $p = 0.018$, $BF_{10} = 8.279$; V-only vs AVbkg, $t_{15} = -0.233$, $p > 0.9$, $BF_{10} = 0.262$). In contrast, no significant amplitude changes of pupil oscillation were found in Experiment 2b where the tone was synchronized with the darker disc ($ts < 1$, $ps > 0.9$; V-only vs AVd: $BF_{10} = 0.337$; AVbkg vs AVd: $BF_{10} = 0.277$; V-only vs AVbkg: $BF_{10} = 0.284$; Fig. 3b). Experiment 2 therefore revealed that audiovisual signals attenuated the strength of pupil oscillation evoked by repeated brighter visual stimuli, while it did not increase the pupil oscillation when the visual stimuli were darker against the background. As hypothesized, the results lend support to the notion that at the relatively fast stimulus repetition speed (e.g., 1.25 Hz), pupillary light reflex can be specifically inhibited in a multisensory context.

According to the principle of inverse effectiveness, the strength of cross-modal stimuli should be relatively low for the largest enhancement of multisensory integration (Noesselt

et al., 2010; Stein & Stanford, 2008; Stein et al., 2020). It may be argued that the failure to reveal an enhanced pupil oscillation in Experiment 2b is attributed to the relative strength rather than relative speed of the induced pupil oscillation. In response to this, we reduced the luminance difference between the visual stimulus and the background and checked whether multisensory signals could enhance pupil oscillation. However, in a supplemental experiment under the same analysis protocol, although repeated presentation of visual stimulus isoluminant with the background induced a pupil oscillation at an extremely low magnitude (~ 0.03 a.u.), we still could not observe an increased pupil oscillation (Supplementary Fig.2).

Furthermore, we noticed that among the four previous studies that reported pupil dilation induced by audiovisual integration, two of them presented stimulus in the peripheral visual field as they were interested in orienting behaviors (Wang et al., 2017; Wang et al., 2014), two of them presented stimulus in the central visual field (Rigato et al., 2016; Van der Stoep et al., 2021). It seems that audiovisual signals are able to dilate pupil size wherever the visual stimulus appears. To further characterize the multisensory modulation of pupil oscillation induced by luminance change, we continued Experiment 3 by moving the visual stimulus from the central to the peripheral field to examine the role the visual eccentricity in the observed effect.

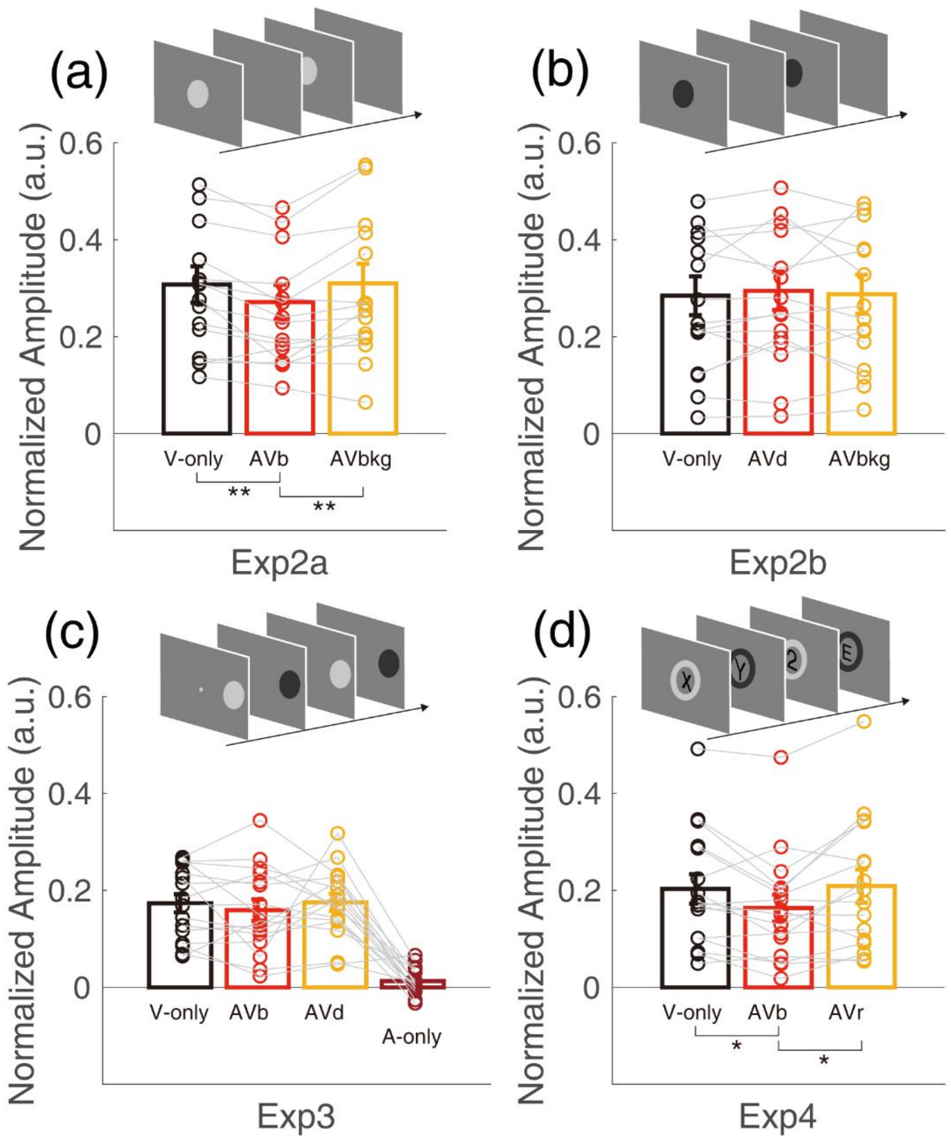

Figure 3. Results of Experiments 2 – 4. The normalized oscillatory amplitude at 1.25 Hz for Experiments 2, 3, and 4, respectively. Each circle represents the amplitude from one participant. The error bar indicates the standard error of mean. ** means $p < 0.01$, * means $p < 0.05$, both uncorrected. AVb represents the AVbright condition; AVd represents the AVdark condition; AVbkg represents the AVbackground condition. AVr represents the AVrandom condition.

4 Experiment 3

Some studies have revealed differential multisensory effects dependent on stimulus eccentricity (Gleiss & Kayser, 2013; Leo et al., 2008; Nidiffer, Stevenson, Fister, Barnett, & Wallace, 2016; van Atteveldt, Peterson, & Schroeder, 2014), but the impact of multisensory integration on pupil size seems to be irrelevant to stimulus eccentricity (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017). Experiment 3 then evaluated whether the audiovisual inhibition of pupillary light reflex remained when the visual stimuli were moved from the central to the peripheral field.

4.1 Methods

4.1.1 Participants

A new group of 16 participants took part in Experiment 3 (10 females; mean age: 23.3 ± 3.9 years).

4.1.2 Stimuli and apparatus

In Experiment 3, the visual stimulus was a disc too, but presented in the left or the right peripheral visual field (eccentricity 10.72 º from the center of the disc to the fixation). The luminance of the disc changed at 1.25 Hz between brightness (47.47 cd/m^2) and darkness (3.03 cd/m^2), as it did in Experiment 1. The luminance range of the disc was expanded because in our preliminary data, the disc had to flicker in a larger luminance range to induce a pupil oscillation whose amplitude may approach that in the central field. The auditory stimulus, still presented binaurally through headphones, but the sound level in the left or right channel was accordingly attenuated 50% to mimic the tones coming from its opposite side. For instance, we would perceive a tone source from the left side, if the sound level of

the right channel is set to be somewhat lower than that of the left channel. Although this manipulation could not precisely align the positions of the tones and flickers, the minor spatial misalignment may not affect the results according to previous findings (Gleiss & Kayser, 2013; van Atteveldt et al., 2014).

4.1.3 Procedures, data analysis and statistics

The procedure, analysis and statistics were all identical to Experiment 1.

4.2 Results and discussion

The accuracies of the oddball counting task were 0.97 ± 0.05 in the V-only condition, 0.98 ± 0.03 in the A-only condition, 0.95 ± 0.07 in the AVb condition, and 0.96 ± 0.04 in the AVd condition. As in Experiments 1 and 2, we observed significant pupil oscillation in the three conditions where the flickering disc was presented, with their amplitudes at 1.25 Hz significantly greater than zero ($ts > 7$, $ps < 2^{e-5}$, $BF_{10} > 6^{e+3}$), but not in the A-only condition ($t_{15} = 1.859$, $p > 0.3$, $BF_{10} = 1.024$; Fig. 3c). However, paired-sample t-tests failed to reveal any significant differences between the amplitudes of pupil oscillation across the three conditions ($ts < 1$, $ps > 0.9$; V-only vs AVb: $BF_{10} = 0.370$; AVd vs AVb: $BF_{10} = 0.322$; V-only vs AVd: $BF_{10} = 0.257$). The evidence is thus prone to support that pupillary light reflex is not inhibited by audiovisual signals when the visual stimulus is presented in the periphery. No inhibition of pupil oscillation in Experiment 3 can neither be attributed to the relatively weaker amplitude of the evoked pupil oscillation (see Fig. 3d), nor be attributed to no audiovisual combination in a repetition paradigm (Noesselt et al., 2007; Talsma & Woldorff, 2005, also see Supplementary Information and Supplementary Fig. 3, where we found the onset pupil size was significantly dilated by audiovisual inputs relative to visual inputs,

consistent with Wang et al., 2017). It is most likely in Experiment 3 that despite being fused, multisensory signals failed to inhibit the pupillary light reflex evoked by a peripheral visual stimuli. This result contrasted with previous findings, which focused on the multisensory impact on the pupil orienting response (Wang et al., 2017; Wang et al., 2014).

So far, the visual flicker was always required to be attended since it was task-relevant. Given several studies have found that even task-irrelevant bimodal signals showed some signs of fusion relative to unimodal signals (Heeman, Nijboer, Van der Stoep, Theeuwes, & Van der Stigchel, 2016; Krause, Schneider, Engel, & Senkowski, 2012; Mühlberg & Müller, 2020; Matusz et al., 2015), it is hypothesized that the inhibition of pupillary light reflex would not be affected despite the visual and auditory stimuli are task-irrelevant and out of attentional focus. We conducted Experiment 4 to test this hypothesis.

5 Experiment 4

Experiment 4 replaced the oddball counting task with a Rapid Stimulus Visual Presentation (RSVP) task following (Santangelo & Spence, 2007) and relocated the visual flicker to the surround of the RSVP so that the visual flicker was now totally task-irrelevant. We examined here whether the induced pupil oscillation was still inhibited when the tone pulses were temporally congruent with the bright phase of the surround visual flicker as in Experiment 1.

5.1 Methods

5.1.1 Participants

Sixteen participants took part in Experiment 4 (9 females; mean age: 22.0 ± 2.3 years).

5.1.2 Stimuli and apparatus

For the visual stimulus, the disc was replaced by a ring (inner circle radius: 1.34º; outer circle radius: 2.68º), with its luminance flickering between 26.8 cd/m^2 and 34.4 cd/m^2 at a frequency of 1.25 Hz. A stream of letters (1.61º ×1.61º) was rapidly presented at 6 Hz within the inner circle of the ring without blank intervals so that each letter lasted 167ms (Fig.3d). The letters were randomly selected from the alphabet, with B, F, I, J, L, O, P, Q, W, and Z excluded. Each letter was always different from its neighbours in the stream. Among the letters, there would embed some numbers of the same size, randomly selected from 2, 3, 4, 6, 7, and 9. The auditory stimulus was identical to Experiment 1.

5.1.3 Procedures

In Experiment 4, participants performed a RSVP task. In each trial, they counted for how many times the numbers appeared (0 – 3 times) among the rapidly presented stream of letters, and were instructed in advance to ignore the flickering ring outside the letter streams during the whole experiment. The visual inducer of the pupil oscillation, therefore, was out of attention focus and should be deemed task-irrelevant. There were 3 conditions, V-only, AVbright, and AVrandom. The V-only and AVb condition were the same as Experiments 1 and 3 except a new AVrandom condition (AVr) was used as a control. In this condition, the tone was not played synchronously with the dark phase of the ring, but randomly played at any possible time from 0.2 secs after the bright phase onset to 0.2 secs before the dark phase offset. Comparison of the pupil oscillations from the AVb and AVr conditions can further demonstrate whether the change of pupillary light reflex is affected by the temporal synchrony between the auditory and visual stimuli. Participants completed a total of 48 trials, divided into 4 blocks, with each condition repeated 16 times.

5.1.4 Data analysis and statistics

The analysis and statistics were same as Experiments 1 – 3.

5.2 Results and discussion

The performance of participants in the oddball counting task was 0.96 ± 0.05 in the V-only condition, 0.97 ± 0.06 in the AVb condition, and 0.93 ± 0.08 in the AVr condition, implying that their attention was concentrated on the RSVP task. Although task-irrelevant, the visual flicker induced significant pupil oscillation as well (Fig.3d, *ts* > 5, *ps* < 1^{e-4}, BF_{10} > 700). The pupil oscillated at a relatively lower amplitude (about 2/3 of the amplitude of Experiments 1 and 2a) probably because the stimuli were unattended (Naber et al., 2013) or eccentrically located. Consistent with Experiments 1 and 2a, the amplitude of pupil oscillation decreased when the tones were temporally congruent with the bright phase of the visual stimuli, compared with when the visual stimuli were alone (V-only vs AVb: t_{15} = 2.904, p = 0.033, BF_{10} = 5.093) and when the audiovisual stimuli were temporally asynchrony (AVr vs AVb: t_{15} = 2.898, p = 0.033, BF_{10} = 5.040; V-only vs AVr: t_{15} = -0.694, p > 0.9, BF_{10} = 0.316). The results indicated that the pupillary light reflex can be inhibited in a multisensory context even though the visual inducer is task-irrelevant. It also demonstrated that the inhibition of pupillary light reflex was sensitive to the cross-modal temporal relationship.

To further explore whether task-relevance modulates such inhibition effect, we calculated an inhibition index (i.e., the difference of oscillatory amplitude between the AVb condition and other conditions, including the V-only, AVd, AVbkg, or AVr conditions, with the latter three conditions represented uniformly by AVincongruent abbreviated as AVinc

for convenience) for Experiments 1, 2a, and 4 separately, then compared the inhibition index of Experiment 4 with those from Experiments 1 and 2a using independent-sample t tests. The results revealed no significant effects [for Experiment 1 vs 4, *ts* < 0.8, *ps* > 0.9, BF_{10} ($Index_{Vonly\text{-}AVb}$) = 0.384, BF_{10} ($Index_{AVinc\text{-}AVb}$) = 0.410; for Experiment 2a vs 4, *ts* < 0.4, *ps* > 0.9, BF_{10} ($Index_{Vonly\text{-}AVb}$) = 0.341, BF_{10} ($Index_{AVinc\text{-}AVb}$) = 0.352]. Taken together, these results are prone to suggest that the inhibition of pupillary light reflex in a multisensory context is immune to task irrelevance.

6 General discussion

Previous studies have shown that multisensory integration enlarges pupil size (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017; Wang et al., 2014). Here using a pupil oscillation frequency tagging method (Naber et al., 2013), the present study for the first time demonstrated that the pupil oscillation evoked by a visual flicker was attenuated when a sequence of tone pulses was synchronized with the bright phase of the visual flicker, relative to when it was synchronized with the dark phase or there was no tones. This implicates that multisensory signals can specifically inhibit the pupillary light reflex when the luminance alternation is at a relatively fast speed (e.g., 1.25 Hz).

As the parasympathetic activation constricts pupil size and the sympathetic activation dilates pupil size (Eckstein et al., 2017; Joshi & Gold, 2020; Larsen & Waters, 2018; Wang & Munoz, 2015), there are parallel explanations for the previously found stronger pupil dilation to multisensory signals (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017; Wang et al., 2014), an inhibited parasympathetic activation, an enhanced sympathetic activation or a combination of them. The currently found inhibition of pupillary

light reflex is likely caused by an inhibition of parasympathetic activation, as the pupillary light reflex is mainly driven by the parasympathetic activation (Clarke, Zhang, & Gamlin, 2003; Joshi & Gold, 2020). But considering the two pupil-related pathways are complicatedly interconnected (ref to Joshi & Gold, 2020, Box 1), the inhibtion of pupillary light reflex may be equally accounted for by an increase of the phasic sympathetic activity, which can dilate pupil size and thereafter counteract the pupillary light reflex. Because both the unimodal and bimodal stimulus were repeated periodically at relatively fast 1.25 Hz in our experiments, only multisensory impact that rapidly changes the trough or the peak of the pupil oscillation within the cyclic period (e.g., 400 ms) could change the amplitude of the pupil oscillation (otherwise the trough and the peak may be equally changed so that the oscillatory amplitude would remain almost the same). The parasympathetic activity, which has a very short onset latency to constrict pupil (< ~270 ms with less than ~800 ms to reach its extreme; Clarke, Zhang, & Gamlin, 2003; Wang & Munoz, 2014), is deemed capable of being transiently inhibited within such limited time. By contrast, the pupil dilation caused by sympathetic activation (primarily through the locus coeruleus-noradrenergic system), which arises slowly with a onset latency of ~330 ms or more (often with a peak latency of more than 1 sec; Chapman, Oka, Bradshaw, Jacobson, & Donaldson, 1999; Liao, Yoneya, Kidani, Kashino, & Furukawa, 2016; Steiner & Barry, 2011; Wang & Munoz, 2014), may be too sluggish to be sufficiently enhanced within this cyclic period. Moreover, we would concurrently observe an enhanced pupil oscillation when the tone synchronized with the dark phase if the phasic sympathetic activation was enhanced. But this was not the case in Experiments 1 and 2.

It might be further argued that this phasic sympathetic activity, albeit arises slowly, may be gradually enhanced and accumulated during repeition of the bimodal inputs, and the inhibited pupillary light reflex may be confounded by the possible pupil dilation caused by this accumulation. Here we provided some further evidence against this possibility. First, although an oddball event can enlarge pupil size, the pupil size for a repeated event would habituate as its novelty gradually decreases (Liao et al., 2016; Netser, Ohayon, & Gutfreund, 2010; Steiner & Barry, 2011). Based on these results, our experiments could hardly lead to a gradual increase of pupil size by periodical presentation of simple emotionally neutral visual stimuli and pure tones. Second, additional analysis, which split the trials into early and late parts, were performed to statistically assess whether the gradual change of pupil size during the stimulus repetition influenced the multisensory inhibition of pupillary light reflex. The analysis for Experiments 1, 2a, and 4 found almost the same results in the early and late parts (and significant inhibition of pupillary light reflex was more frequently found in the early part), which indicated little evidence for gradual pupil dilation and corresponding confounding on our main observation (for detailed analysis, see Supplementary Information).

Of note, although it is more likely the inhibition of parasympathetic activation that accounts for our observation, we do not claim that the sympathetic activation cannot be enhanced in a multisensory context. Dissimilar to the parasympathetic pathway that can be transiently inhibited, we propose the sympathetic pathway may be enhanced by multisensory signals in a slow and sustained manner. This is compatible with previous findings, which demonstrated that the pupil dilation to multisensory signals could on one

hand be as early as that of the pupillary light reflex (Wang et al., 2017), while on the other hand arise late and sustain for a relatively long time (Rigato et al., 2016; Van der Stoep et al., 2021). This assumption can also explain the inconsistency between our observation and a recent one (Van der Stoep et al., 2021), which reported no distinction between phasic pupil response to light and dark with each trial only including one unimodal or bimodal stimuli but with adequate time to observe the pupil change.

Put aside the possible explanations about the underpinning pathway, the present study further revealed that the multisensory inhibition of pupillary light reflex can only be observed when the visual flicker was located at the central field. The result is in contrast with the findings that pupil dilation by multisensory signals may be independent of stimulus eccentricity (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017; Wang et al., 2014). But it is not completely unexpected since multisensory integration in the central and peripheral fields has been proposed to be functionally complementary. Stimuli in the central field may be prioritized in accurate discrimination and recognition with regard to their properties and features, whereas stimuli in the periphery may signal potential threat, which require fast orienting response either in an overt or covert manner (Chen, Maurer, Lewis, Spence, & Shore, 2017; Gleiss & Kayser, 2013; Leo et al., 2008; Nidiffer et al., 2016; van Atteveldt et al., 2014). It is thus possible that once the visual flicker had already attracted covert attention in Experiment 3 which required to fixate at the center, overt orienting responses, such as to saccade towards the target location, would be suppressed thereafter. Given that the superior colliculus (SC) is an important nucleus for both saccade generation (Coe & Munoz, 2017) and multisensory integration (King, 2004; Stein & Stanford, 2008;

Stein et al., 2020), suppression of saccades may be accompanied by an attenuation of multisensory interaction in SC. This probably leads to no multisensory modulation of pupillary light reflex in the periphery.

Although dependent on stimulus eccentricity, fusion of multisensory inputs has been proposed independent of task relevance. Previous studies have reported that even task-irrelevant cross-modal signals can exert a stronger interference on the currently performed task compared to a unimodal distractor (Heeman et al., 2016; Krause et al., 2012; Matusz et al., 2015, but an improvement in Mühlberg & Müller, 2020 and no effect in Experiment 4 of Lunn, Sjoblom, Ward, Soto-Faraco, & Forster, 2019). Despite that no interference on the RSVP task was found in the present study, the pupillary light reflex induced by the visual stimuli that were task-irrelevant and out of attentional focus was inhibited by temporally congruent tone pulses in Experiment 4. The result verified that the multisensory inhibition of pupillary light reflex may be insensitive to the attentional set defined by the goal, and perhaps controlled by a bottom-up, stimulus-driven mechanism. Moreover, it suggests the changes of pupil size can be an effective physiological proxy for a task-irrelevant multisensory effect, similar to other index, for instance, the steady state visual evoked potentials (Krause et al., 2012). But notably, task irrelevance does not necessarily mean immunity to attentional load. The higher RSVP accuracy in Experiment 4 ensures the task-relevant stimuli being fully attended on one hand, but indicates an attentional load perhaps at a medium level on anther hand. As several studies reported that the effect of multisensory integration would be attenuated at higher attentional load (Fairhall & Macaluso, 2009; Morís Fernández, Visser, Ventura-Campos, Ávila, & Soto-Faraco, 2015;

Senkowski, Talsma, Herrmann, & Woldorff, 2005; Talsma, Doty, & Woldorff, 2007; Talsma & Woldorff, 2005, but see Santangelo & Spence, 2007; Wahn & König, 2015), it remains to be sought out in the future how the pupillary light reflex in a multisensory context would be when the attentional load is strongly increased.

Regarding to the neural node related to this multisensory influence of pupillary light reflex, we infer that the most relevant structure is SC. SC has been shown to project directly or indirectly to the pretectal olivary nucleus and the Edinger-Westphal nucleus on the parasympathetic pathway (Harting, Huerta, Frankfurter, Strominger, & Royce, 1980; May, 2006; May, Warren, Bohlen, Barnerssoi, & Horn, 2016; Wang & Munoz, 2015). It also receives input from LC and may indirectly influence the sympathetic pathway through the mesencephalic cuneiform nucleus (Joshi & Gold, 2020; Wang & Munoz, 2015). Electrical microstimulation of the intermediate layers of SC could produce transiently pupil dilation, verifying the ability of SC in modulating pupil size (Wang et al., 2014; Wang, Boehnke, White, & Munoz, 2012). Importantly, SC whose deeper layers are able to integrate multisensory inputs, is repeatedly proved to be a subcortical hub of multisensory integration (Stein & Stanford, 2008; Stein et al., 2020). Taken together, it is most probable that SC first combines the temporally congruent auditory and visual inputs, and then modulates the pupil size through suppressing the parasympathetic activation (or enhancing the sympathetic activation). The cross-modal integration in SC is also compatible with the observed stimulus eccentricity dependence, as discussed earlier. But it still remains possible that the auditory inputs may directly inhibit the parasympathetic activity (or increase the sympathetic activity) through LC (Joshi, Li, Kalwani, & Gold, 2016). It is hard

to disentangle how multisensory signals are interacted to affect the pupillary light reflex purely from the physiological data reported here, although SC might be a key neural candidate involved in this process.

In conclusion, the present study demonstrated that pupillary light reflex in response to a central visual inducer can be specifically inhibited in a multisensory context regardless of task relevance. This inhibition of pupillary light reflex not only implies the capability of multisensory signals to mediate the pupil-related neural pathway, but also provides another easily measured pupillometric indicator of multisensory interaction independent of explicit response. Intriguingly, if there are signals from other modalities capable of promoting pupil constriction, would an increased pupillary light reflex be specifically observed? This would be regarded as a complementary to the current findings.

Data Accessibility

All the data used for statistics, and code to generate the figures could be found using https://osf.io/npaer/?view_only=287f4f90a4304065b4aecf243246f134.

Competing Interests

The authors declared no competing financial interest.

Funding

The study was supported by research grants from the National Natural Science Foundation of China (31830037, 31525011, 31600884), the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (Y8CX372005), and the Strategic Priority Research Program and the Key Research Program of Frontier Sciences of the Chinese

Academy of Sciences (XDB32010300, QYZDB-SSW-SMC030), Beijing Municipal Science & Technology Commission, Shenzhen-Hong Kong Institute of Brain Science, and Fundamental Research Funds for the Central Universities.

Acknowledgment

We appreciate the critical comments of Dr. Ying Wang and Dr. Xilei Zhang on the study.

References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433. doi: 10.1163/156856897X00357

Chapman, C. R., Oka, S., Bradshaw, D. H., Jacobson, R. C., & Donaldson, G. W. (1999). Phasic pupil dilation response to noxious stimulation in normal volunteers: Relationship to brain evoked potentials and pain report. *Psychophysiology*, *36*(1), 44-52. doi: https://doi.org/10.1017/S0048577299970373

Chen, Y.-C., Maurer, D., Lewis, T. L., Spence, C., & Shore, D. I. (2017). Central–peripheral differences in audiovisual and visuotactile event perception. *Attention, Perception, & Psychophysics, 79*(8), 2552-2563. doi: 10.3758/s13414-017-1396-4

Clarke, R. J., Zhang, H., & Gamlin, P. D. R. (2003). Characteristics of the Pupillary Light Reflex in the Alert Rhesus Monkey. *Journal of neurophysiology*, *89*(6), 3179-3189. doi: 10.1152/jn.01131.2002

Coe, B. C., & Munoz, D. P. (2017). Mechanisms of saccade suppression revealed in the anti-saccade task. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1718), 20160192. doi: doi:10.1098/rstb.2016.0192

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology, 5*(781). doi: 10.3389/fpsyg.2014.00781

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience, 25*, 69-91. doi: 10.1016/j.dcn.2016.11.001

Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at

multiple cortical and subcortical sites. *European Journal of Neuroscience, 29*(6), 1247-1257. doi: 10.1111/j.1460-9568.2009.06688.x

Gleiss, S., & Kayser, C. (2013). Eccentricity dependent auditory enhancement of visual stimulus detection but not discrimination. *Frontiers in Integrative Neuroscience, 7*, 52. doi: 10.3389/fnint.2013.00052

Harting, J. K., Huerta, M. F., Frankfurter, A. J., Strominger, N. L., & Royce, G. J. (1980). Ascending pathways from the monkey superior colliculus: An autoradiographic analysis. *Journal of Comparative Neurology, 192*(4), 853-882. doi: 10.1002/cne.901920414

Heeman, J., Nijboer, T. C. W., Van der Stoep, N., Theeuwes, J., & Van der Stigchel, S. (2016). Oculomotor interference of bimodal distractors. *Vision Research, 123*, 46-55. doi: 10.1016/j.visres.2016.04.002

Joshi, S., & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. *Trends in Cognitive Sciences, 24*(6), 466-480. doi: 10.1016/j.tics.2020.03.005

Joshi, S., Li, Y., Kalwani, Rishi M., & Gold, Joshua I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron, 89*(1), 221-234. doi: 10.1016/j.neuron.2015.11.028

Kayser, S. J., Philiastides, M. G., & Kayser, C. (2017). Sounds facilitate visual motion discrimination via the enhancement of late occipital visual representations. *NeuroImage, 148*, 31-41. doi: 10.1016/j.neuroimage.2017.01.010

King, A. J. (2004). The superior colliculus. *Current Biology, 14*(9), R335-R338. doi: https://doi.org/10.1016/j.cub.2004.04.018

Krause, H., Schneider, T., Engel, A., & Senkowski, D. (2012). Capture of visual attention

interferes with multisensory speech processing. *Frontiers in Integrative Neuroscience, 6*(67). doi: 10.3389/fnint.2012.00067

Larsen, R. S., & Waters, J. (2018). Neuromodulatory Correlates of Pupil Dilation. *Frontiers in Neural Circuits, 12*(21). doi: 10.3389/fncir.2018.00021

Leo, F., Bertini, C., di Pellegrino, G., & Làdavas, E. (2008). Multisensory integration for orienting responses in humans requires the activation of the superior colliculus. *Experimental Brain Research, 186*(1), 67-77. doi: 10.1007/s00221-007-1204-9

Lewis, R., & Noppeney, U. (2010). Audiovisual Synchrony Improves Motion Discrimination via Enhanced Connectivity between Early Visual and Auditory Areas. *The Journal of Neuroscience, 30*(37), 12329-12339. doi: 10.1523/jneurosci.5745-09.2010

Liao, H.-I., Kidani, S., Yoneya, M., Kashino, M., & Furukawa, S. (2016). Correspondences among pupillary dilation response, subjective salience of sounds, and loudness. *Psychonomic Bulletin & Review, 23*(2), 412-425. doi: 10.3758/s13423-015-0898-0

Liao, H.-I., Yoneya, M., Kidani, S., Kashino, M., & Furukawa, S. (2016). Human Pupillary Dilation Response to Deviant Auditory Stimuli: Effects of Stimulus Properties and Voluntary Attention. *Frontiers in Neuroscience, 10*(43). doi: 10.3389/fnins.2016.00043

Lunn, J., Sjoblom, A., Ward, J., Soto-Faraco, S., & Forster, S. (2019). Multisensory enhancement of attention depends on whether you are already paying attention. *Cognition, 187*, 38-49. doi: https://doi.org/10.1016/j.cognition.2019.02.008

Mühlberg, S., & Müller, M. M. (2020). Alignment of Continuous Auditory and Visual Distractor Stimuli Is Leading to an Increased Performance. *Frontiers in psychology, 11*(790). doi: 10.3389/fpsyg.2020.00790

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition, 136*, 156-165. doi: 10.1016/j.cognition.2014.11.031

May, P. J. (2006). The mammalian superior colliculus: laminar structure and connections. In J. A. Büttner-Ennever (Ed.), *Progress in Brain Research* (Vol. 151, pp. 321-378): Elsevier.

May, P. J., Warren, S., Bohlen, M. O., Barnerssoi, M., & Horn, A. K. E. (2016). A central mesencephalic reticular formation projection to the Edinger–Westphal nuclei. *Brain Structure and Function, 221*(8), 4073-4089. doi: 10.1007/s00429-015-1147-z

Morís Fernández, L., Visser, M., Ventura-Campos, N., Ávila, C., & Soto-Faraco, S. (2015). Top-down attention regulates the neural expression of audiovisual integration. *NeuroImage, 119*, 272-285. doi: 10.1016/j.neuroimage.2015.06.052

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology, 4*. doi: 10.3389/fpsyg.2013.00919

Netser, S., Ohayon, S., & Gutfreund, Y. (2010). Multiple Manifestations of Microstimulation in the Optic Tectum: Eye Movements, Pupil Dilations, and Sensory Priming. *Journal of neurophysiology, 104*(1), 108-118. doi: 10.1152/jn.01142.2009

Nidiffer, A. R., Stevenson, R. A., Fister, J. K., Barnett, Z. P., & Wallace, M. T. (2016). Interactions between space and effectiveness in Human multisensory performance. *Neuropsychologia, 88*, 83-91. doi: 10.1016/j.neuropsychologia.2016.01.031

Noesselt, T., Bergmann, D., Hake, M., Heinze, H.-J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Research, 1220*, 157-163.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., & Driver, J. (2007). Audiovisual Temporal Correspondence Modulates Human Multisensory Superior Temporal Sulcus Plus Primary Sensory Cortices. *Journal of Neuroscience., 27*(42), 11431-11441. doi: 10.1523/jneurosci.2252-07.2007

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H.-J., & Driver, J. (2010). Sound-Induced Enhancement of Low-Intensity Vision: Multisensory Influences on Human Sensory-Specific Cortices and Thalamic Bodies Relate to Perceptual Enhancement of Visual Detection Sensitivity. *J. Neurosci., 30*(41), 13609-13623. doi: 10.1523/jneurosci.4524-09.2010

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis, 10*(4), 437-442. doi: https://doi.org/10.1163/156856897X00366

Rigato, S., Rieger, G., & Romei, V. (2016). Multisensory signalling enhances pupil dilation. *Scientific Reports, 6*, 26188. doi: 10.1038/srep26188

Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of experimental psychology: Human perception and performance, 33*(6), 1311-1321. doi: 10.1037/0096-1523.33.6.1311

Senkowski, D., Talsma, D., Herrmann, C., & Woldorff, M. (2005). Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Experimental Brain Research, 166*(3), 411-426. doi: 10.1007/s00221-005-2381-z

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci, 9*(4), 255-266.

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2020). Multisensory Integration and the Society

for Neuroscience: Then and Now. *The Journal of Neuroscience, 40*(1), 3-11. doi: 10.1523/jneurosci.0737-19.2019

Steiner, G. Z., & Barry, R. J. (2011). Pupillary responses and event-related potentials as indices of the orienting reflex. *Psychophysiology, 48*(12), 1648-1655. doi: 10.1111/j.1469-8986.2011.01271.x

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? *Cerebral Cortex, 17*(3), 679-690. doi: 10.1093/cercor/bhk016

Talsma, D., & Woldorff, M. G. (2005). Selective Attention and Multisensory Integration: Multiple Phases of Effects on the Evoked Brain Activity. *Journal of cognitive neuroscience, 17*(7), 1098-1114. doi: doi:10.1162/0898929054475172

van Atteveldt, N. M., Peterson, B. S., & Schroeder, C. E. (2014). Contextual control of audiovisual integration in low-level sensory cortices. *Human Brain Mapping, 35*(5), 2394-2411. doi: 10.1002/hbm.22336

Van der Burg, E., Olivers, C., Bronkhorst, A., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of experimental psychology. Human perception and performance, 34*(5), 1053-1065.

Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage, 55*(3), 1208-1218. doi: 10.1016/j.neuroimage.2010.12.068

Van der Stoep, N., Van der Smagt, M. J., Notaro, C., Spock, Z., & Naber, M. (2021). The additive nature of the human multisensory evoked pupil response. Scientific Reports, 11(1), 707.

doi: 10.1038/s41598-020-80286-1

Wahn, B., & König, P. (2015). Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. *Frontiers in psychology, 6*. doi: 10.3389/fpsyg.2015.01084

Wang, C.-A., Blohm, G., Huang, J., Boehnke, S. E., & Munoz, D. P. (2017). Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. *Biological Psychology, 129*(Supplement C), 36-44. doi: 10.1016/j.biopsycho.2017.07.024

Wang, C.-A., Boehnke, S. E., Itti, L., & Munoz, D. P. (2014). Transient Pupil Response Is Modulated by Contrast-Based Saliency. *The Journal of Neuroscience, 34*(2), 408-417. doi: 10.1523/jneurosci.3550-13.2014

Wang, C.-A., Boehnke, S. E., White, B. J., & Munoz, D. P. (2012). Microstimulation of the Monkey Superior Colliculus Induces Pupil Dilation Without Evoking Saccades. *The Journal of Neuroscience, 32*(11), 3629-3636. doi: 10.1523/jneurosci.5512-11.2012

Wang, C.-A., & Munoz, D. P. (2014). Modulation of stimulus contrast on the human pupil orienting response. *European Journal of Neuroscience, 40*(5), 2822-2832. doi: 10.1111/ejn.12641

Wang, C.-A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology, 33*, 134-140. doi: 10.1016/j.conb.2015.03.018

Werner, S., & Noppeney, U. (2010). Distinct Functional Contributions of Primary Sensory and Association Areas to Audiovisual Integration in Object Categorization. *The Journal of Neuroscience, 30*(7), 2662-2675. doi: 10.1523/jneurosci.5091-09.2010

Werner, S., & Noppeney, U. (2011). The Contributions of Transient and Sustained Response Codes to Audiovisual Integration. *Cerebral Cortex, 21*(4), 920-931. doi: 10.1093/cercor/bhq161